JOURNAL OF VIROLOGY, Jan. 1987, p. 66–74
0022-538X/87/010066-09$02.00/0

Vol. 61, No. 1

Mammary Tumorigenesis in Feral Mice: Identification of a New *int* Locus in Mouse Mammary Tumor Virus (Czech II)-Induced Mammary Tumors

DANIEL GALLAHAN AND ROBERT CALLAHAN*

Laboratory of Tumor Immunology and Biology, National Cancer Institute, Bethesda, Maryland 20892

Received 11 August 1986/Accepted 9 October 1986

A population of *Mus musculus* subsp. *musculus* (Czech II), recently isolated from the wild, lack endogenous mouse mammary tumor virus (MMTV) proviral genomes. Some of these mice carry an infectious MMTV [designated MMTV (Czech II)] that is transmitted in the milk and is associated with mammary tumor development. This virus is distinct from laboratory strains of MMTV present in inbred mice. An MMTV (Czech II) genome was found within a 0.5-kilobase region of the cellular genome in five of 16 Czech II mammary tumors. MMTV insertion at this site activates expression of a 2.4-kilobase species of RNA from a previously silent cellular gene. This region of the cellular genome was designated *int-3* since it is unrelated to the *int-1* and *int-2* loci. The *int-3* locus does not appear to correspond to other proto-oncogenes but is well conserved among mammalian species.

Spontaneous mammary tumorigenesis in mice is frequently associated with chronic infection of host mammary tissue by mouse mammary tumor virus (MMTV) (38). MMTV, like other retroviruses, can act as an insertional mutagen since it can integrate at numerous, perhaps, random, sites in the cellular genome of the host (36). It has been proposed as a general hypothesis that cellular genes, whose expression is activated or augmented as a result of retroviral insertion in many tumors of the same type, probably contribute to tumorigenesis (36). In avian leukosis virus-induced B-cell lymphomas the activated cellular gene is c-*myc*, the progenitor of a retroviral oncogene (15). The situation in MMTV-induced mammary tumors appears more complex. The association between MMTV and mammary tumorigenesis was first recognized in inbred strains of *Mus musculus* which had been selectively bred for a high incidence of mammary tumors (37). Analysis of mammary tumors from the C3H and BR6 mouse strains led to the identification of two cellular genetic loci (designated *int-1* and *int-2*) which are frequently occupied by an MMTV genome in mammary tumor DNA (21, 22). The *int-1* and *int-2* loci are located, respectively, on mouse chromosomes 15 and 7 (20, 23). Insertion of a viral genome at either locus activates expression of a previously silent gene within the affected locus (8, 21). These activated cellular genes are highly conserved among vertebrates and are unrelated to each other or to the known proto-oncogenes (4, 17, 20, 35). The frequent insertion of an MMTV genome at these loci is not limited to mammary tumors of high incidence in inbred mouse strains. Indeed, they are similarly affected in mammary tumors of the distantly related feral species *M. cervicolor* subsp. *popaeus* (10). In this species the MMTV (designated MC-MTV) genome shares limited nucleotide sequence homology with those of laboratory strains of MMTV (18, 28). Although direct evidence for the oncogenic potential of the *int* loci is not available, their activation by different strains of MMTV in distantly related species further supports the hypothesis that they are members of a new family of tumor-associated genes.

* Corresponding author.

Two observations suggest that there might be additional members in this family of tumor-associated genes. (i) Not all MMTV-induced mammary tumors in the C3H, BR6, and *M. cervicolor* subsp. *popaeus* strains (10, 10, and 50%, respectively) contain a mutated *int-1* or *int-2* locus (10, 21, 22, 25). (ii) The frequency with which the *int-1* and *int-2* loci are mutated appears to depend on the strain of mice or MMTV. Thus, the *int-1* locus is primarily affected in C3H and *M. cervicolor* subsp. *popaeus* mammary tumors (80 and 45%, respectively), whereas the *int-2* locus is affected in 70% of BR6 mice and to a lesser extent in the other two strains (10 and 5%, respectively). Although the basis for this apparent specificity is unknown, we reasoned that examination of other MMTV-mouse strain combinations might lead to identification of new *int* loci. To approach this question, we studied mammary tumorigenesis in pedigreed outbred colonies of feral species of the genus *Mus*. One colony (designated Czech II) was derived from a single breeding pair of *M. musculus* subsp. *musculus* trapped in Czechoslovakia. Czech II mice lack endogenous MMTV genomes, but contain a poorly infectious strain of MMTV which is transmitted congenitally by milk (3; D. Gallahan, Ph.D. dissertation, University of Maryland, College Park, 1986). This colony has a 12% incidence of pregnancy-independent mammary adenocarcinomas which are histopathologically similar to those induced by MMTV (C3H). We describe here a new *int* locus (designated *int-3*) in MMTV (Czech II)-induced mammary tumors which is unrelated to *int-1* and *int-2*.

MATERIALS AND METHODS

Source of tumors and tissue. *M. musculus* subsp. *musculus* (Czech II) mice were originally trapped in Sladeckovce, Czechoslovakia, and established in the laboratory from a single breeding pair (3). All mice used were from breeding stocks maintained by the National Cancer Institute at Hazelton Laboratories of America, Inc., Rockville, Md., under National Cancer Institute contract N01-CB-94326. Mice were kept under controlled photoperiods and temperature conditions. Food and water were given ad libitum. Feral mice also received additional food supplements in the form of mealworm larvae. Breeding females were examined

weekly for spontaneous mammary tumors. Suspected abnormal growth in mammary tissue was excised from the mice for histological typing and nucleic acid extraction. Tissue for nucleic acid extraction was frozen in liquid nitrogen and stored at −80°C.

Isolation of high-molecular-weight cellular DNA. Frozen tissue was homogenized in a Waring blender containing an equal weight per volume of 10 mM Tris (pH 8.0) and 5.0 mM EDTA (10 ml/g of tissue). The homogenate was transferred to a 50-ml tube which contained 0.5% sodium dodecyl sulfate, 0.1 M β-mercaptoethanol, and 50 μg of proteinase K (Boehringer Mannhein Biochemicals) per ml. The mixture was then incubated at 37°C for approximately 4 h. DNA was extracted with a mixture of phenol and chloroform as previously described (14). After treatment with RNase, high-molecular-weight DNA was precipitated with ethanol and spooled onto a glass rod. Partially dried DNA was suspended in TE buffer (10 mM Tris [pH 8.0], 1 mM EDTA).

Restriction enzyme digestion of DNA and analysis. Restriction enzyme reactions were carried out as directed by the manufacturer. Restricted DNA samples were electrophoresed on horizontal 0.8% agarose gels in E buffer (40 mM Tris [pH 8.3], 2 mM disodium EDTA) at a constant voltage (50 V) for approximately 18 h. A lane of *Hin*dIII-digested bacteriophage lambda DNA was included as a molecular weight control. The DNA was transferred from the agarose gel to Genetran (Plasco Inc.) nitrocellulose paper for subsequent hybridization (30).

Before hybridization, filters were soaked for 4 to 24 h at 37°C in a prehybridization solution containing 3× SSPE (1× SSPE is 180 mM NaCl, 10 mM NaH_2PO_4 [pH 7.4], and 1 mM EDTA), 5× Denhardt solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin) (7), 2.5% dextran sulfate, and 50% formamide. Denatured probe was added to a solution similar to the prehybridization solution except that it contained 40% formamide. The mixture was added to the plastic bag containing prehybridized filters and incubated at 37°C for 24 h. The filters were washed in stringent conditions involving three changes of 0.1× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 0.5% sodium dodecyl sulfate at 65°C. The filters were exposed to Kodak XAR-5 C-ray film overnight or for up to several days.

Re-use of Genetran membranes containing digested DNA. Genetran filters were stripped of probe DNA and re-used for subsequent hybridization with different probes. The filters were placed in 100 ml of 0.4 N NaOH and incubated for 20 min at room temperature. This solution was discarded and replaced with 0.1 M Tris (pH 7.5)–0.1× SSC–0.5% sodium dodecyl sulfate solution to neutralize the filter. The filter was incubated for 15 min at room temperature, after which the neutralization solution was replaced and incubation was continued for another 15 min. The filter was then prehybridized for subsequent hybridization. Genetran filters treated in this way could be re-used at least five times before noticeable DNA loss.

Isolation of RNA and analysis. Frozen mouse tissues were homogenized in 5 volumes of RNA extraction buffer (6 M guanidinium thiocyanate, 5 mM sodium citrate [pH 7.0], 0.1 M β-mercaptoethanol, 0.5% *N*-laurylsarcosine). The homogenate was layered onto a 1.2-ml cushion of 5.7 M CsCl in 0.1 M EDTA and then centrifuged at 200,000 × *g* for 12 h at 20°C. After centrifugation, the supernatant was discarded, and the pellet was washed in 0°C 70% ethanol and allowed to air dry. The RNA pellet was suspended in a small volume of RNA suspension buffer (10 mM Tris [pH 7.4], 5 mM EDTA, 1% sodium dodecyl sulfate). A 4:1 mixture of chloroform-butanol was used to extract the aqueous RNA-containing phase. The RNA was precipitated with ethanol and suspended in H_2O. Poly(A)$^+$ RNA was selected by oligo(dT)-cellulose columns as previously described (1).

RNA samples were electrophoresed on horizontal gels containing 1.5% agarose in 1× RNA running buffer (20 mM MOPS [morpholinepropanesulfonic acid], 5 mM sodium acetate, 5 mM EDTA [pH 7.0]) and 18% formaldehyde. Poly(A)-selected RNA (10 μg) was precipitated and dried in a 1.5-ml Eppendorf tube. The RNA was dissolved in sample buffer (20 μl) containing 60% formamide, 1.2× running buffer, and 0.2% bromophenol blue. The gels were run at 40 to 50 V overnight with recirculating buffer and then blotted onto a nitrocellulose membrane (33).

The RNA filter was prehybridized, hybridized, and washed under identical conditions for DNA hybridization. After washing, the filters were dried and exposed to X-ray film as with the DNA filter.

Preparation of DNA probes. Plasmids containing the MMTV (C3H) long terminal repeat (LTR) and *gag*, *pol*, and *env* genes were generously provided by H. Varmus, J. Majors, and N. Hynes and have been described elsewhere (3, 27). They are, briefly, LTR, a 1.4-kilobase-pair (kbp) *Pst*I fragment from MMTV (C3H); *gag*, a 2.0-kbp *Pst*I-*Xho*I fragment from MMTV GR-40; *pol*, a 2.4-kbp *Xho*I-*Eco*RI fragment from MMTV GR-40; and *env*, a 1.7-kbp *Pst*I fragment from MMTV (C3H). The *int-1* 2.4-kbp *Eco*RI-*Bam*HI fragment, as well as the 2.0-kbp *Sac*I and 1.8-kbp *Sac*I-*Eco*RI fragments from the 5′ and 3′ ends, respectively, of the *int-2* locus, was generously provided by R. Nusse and C. Dickson. Recombinant plasmids containing viral oncogenes or proto-oncogenes were generously provided by C. Sherr (v-*fes* and v-*fms*), R. Gallo (c-*sis*), E. Chang (human c-Ha-*ras*-1), and J. Featherston (β-actin) or obtained from the American Type Culture Collection (v-*abl*, v-*src*, N-*myc*, N-*ras*, c-*mos*, and cK-*ras*-2). The DNA probes were labeled with [^{32}P]dCTP by nick translation to a specific activity of 10^8 cpm/μg (26).

Molecular cloning of viral and host flanking sequences. The majority of techniques used for isolation and purification of molecular clones were as described by Maniatis et al. (16). Mouse liver or tumor DNA was digested with *Eco*RI in preparative amounts. Specific fragments of interest were identified by hybridization and purified by electrophoresis. These fragments were ligated into λ g^+ Wes λ B. A genomic library was constructed from tumor B4625 DNA partially digested with *Mbo*I. Approximately 15-kilobase (kb) fragments were purified and ligated into *Bam*HI-digested EMBL3 (Vector Cloning Systems). After packaging, the recombinant phage were plated on *Escherichia coli* LE392 and screened by previously described hybridization conditions (16).

RESULTS

Spontaneous mammary tumors in Czech II mice and MMTV (Czech II). Czech II mice have been successfully bred in the laboratory since 1975 (3). While the mice were being maintained, they were closely monitored for disease. Since occasional spontaneous mammary tumors were observed in breeding Czech II females, lactating Czech II mice were tested for MMTV antigens. MMTV particle-associated gp36 was detected in the milk of some females, suggesting the presence of infectious virus in some but not all mice (Gallahan, Ph.D. dissertation). This was confirmed by the

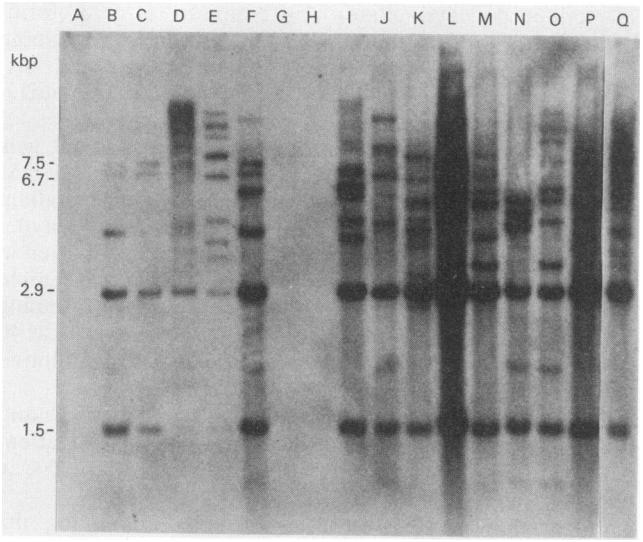

FIG. 1. The presence of integrated MMTV (Czech II) proviral genomes in *M. musculus* subsp. *musculus* (Czech II) mammary tumor DNA compared with normal liver DNA. Mammary tumor and normal liver DNAs (10 μg each) were digested with *Eco*RI, run on a 0.8% agarose gel, and transferred to a nylon membrane. The cellular DNAs were from (lanes) normal liver (A) and mammary tumors B4625 (B), N402 (C), B4501 (D), B4532 (E), B4571 (F), B4579 (G), B4651 (H), B4661 (I), B4603 (J), B4569 (K), B4681 (L), B4559 (M), B4603 no. 2 (N), B4637 (O), B4817 (P), and B4879 no. 2 (Q). The membranes were hybridized with a mixture of the MMTV (strain GR) and LTR probes (see the text).

presence of MMTV RNA in some lactating mammary gland and mammary tumor tissue (Gallahan, Ph.D. dissertation). The virus was designated MMTV (Czech II). The first mammary tumors were identified in the general colony at generation 5. Tumors arose in mice that were active breeders and were classified as pregnancy-independent type A adenocarcinomas (9). These tumors showed the usual well-differentiated alveolar morphology with characteristic acinar structures (data not shown). The Czech II colony had an overall tumor incidence of approximately 12%, although this varied within different maternal lineages. Tumor latency also varied from 9 to 28 months, with an average of 15 months.

Presence of MMTV in Czech II mammary tumors. Czech II mammary tumor DNA was examined for the presence of integrated MMTV (Czech II) proviral genomes. Figure 1 shows blot hybridization results of *Eco*RI-restricted DNAs from 16 Czech II mammary tumors. A mixture of probes representing the entire MMTV genome was used in this study (see Materials and Methods). Tumor DNAs B4579 and B4651 apparently contain no MMTV proviral genomes. This is consistent with the absence of virus in the milk of the mice (data not shown). The other mammary tumor DNAs contained a varied number of MMTV hybridizing fragments. Thus, 14 (87%) of the 16 Czech II mammary tumors examined contained integrated MMTV proviral genomes. In addition to the Czech II liver DNA shown in Fig. 1, a random sampling of Czech II individuals failed to produce evidence for endogenous MMTV proviral genomes within normal tissue (data not shown).

More information about the nature of the integrated viral genome was obtained by digesting DNA from tumor B4625 with restriction enzymes *Pst*I and *Eco*RI. The blots were sequentially hybridized with probes representing the MMTV LTR and *gag*, *pol*, and *env* genes (see Materials and Methods). The results of these hybridizations are shown in Fig. 2A. Since the 2.9- and 1.5-kb *Eco*RI fragments contained, respectively, *gag-pol* and *env* sequences, they represent internal fragments of the integrated proviral genome. These fragments appear to result from the presence of three internal *Eco*RI sites. The presence of a 6.0-kbp *Pst*I fragment which reacted with the *gag*, *pol*, and *env* probes suggested that the MMTV (Czech II) genome lacks a *Pst*I site in the *env* gene which is common in the proviral genomes of laboratory strains of MMTV. The 7.5-kb *Eco*RI fragment which hybridizes to LTR and *env* represents the 3′ end of an integrated provirus and 3′ host flanking sequence. Similarly,

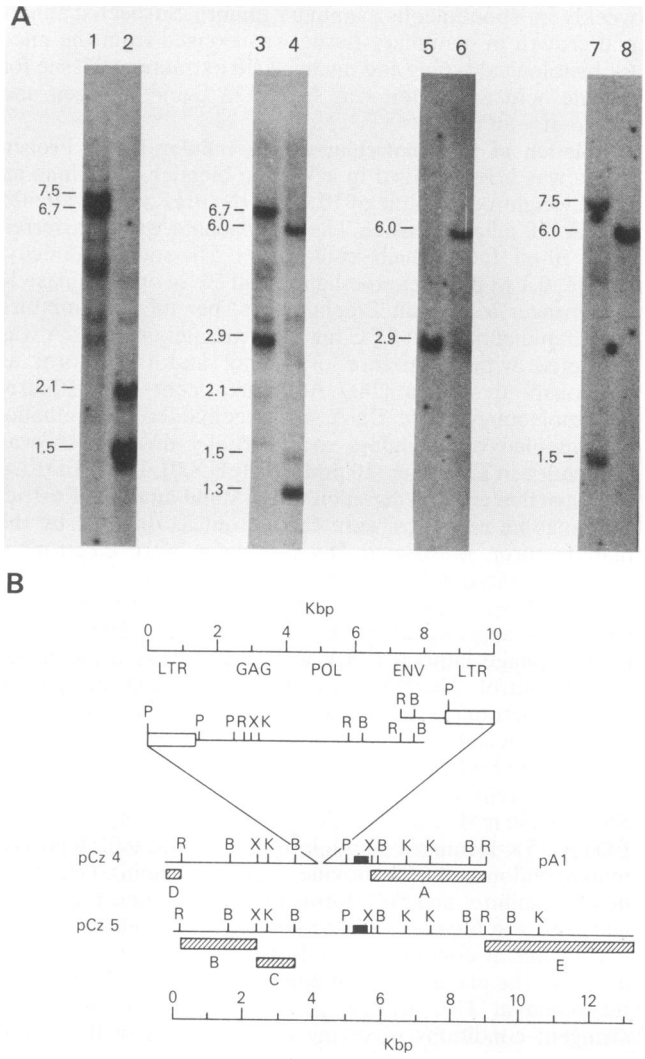

FIG. 2. Organization of MMTV (Czech II) proviral DNA in tumor DNA B4625. (A) Cellular DNA (10 μg) from tumor B4625 was digested with *Eco*RI (lanes 1, 3, 5, and 7) and *Pst*I (lanes 2, 4, 6, and 8). The restricted DNAs were run on the same 0.8% agarose gel and transferred to a nylon membrane. Membranes were hybridized with MMTV LTR (lanes 1 and 2), MMTV *gag* (lanes 3 and 4), MMTV *pol* (lanes 5 and 6), and MMTV *env* (lanes 7 and 8) probes (see the text). (B) Partial restriction maps of recombinant clones pA1, pCz4, and pCz5 (see the text). The restriction sites are indicated as follows: E, *Eco*RI; B, *Bam*HI; P, *Pst*I; X, *Xba*I; K, *Kpn*I. The repetitive cellular DNA sequences are indicated by a solid box. The MMTV LTR is represented by an open box. The unique cellular flanking restriction fragments used as probes, and their origins, are indicated by hatched boxes.

the LTF-*gag*-related 6.7-kb *Eco*RI fragment represents the 5′ end of the virus with host flanking sequences. The 7.5- and 6.7-kb *Eco*RI fragments were present in at least three tumor DNAs (N402, B4625, and B4571) and appear to represent insertions of the viral genome into a common region of the cellular DNA (Fig. 1). Two common integration regions (designated *int-1* and *int-2*) have previously been shown to be frequently occupied by an MMTV proviral genome in mammary tumor DNA (21, 22). However, examination of these genetic loci in all 16 Czech II tumor DNAs demonstrated that they were unaltered (data not shown).

Cloning of the provirus-host junction DNA and the corresponding uninterrupted cellular DNA. These results suggested to us that some of the Czech II tumor DNAs might contain a new common integration region. To explore this possibility and to further characterize the MMTV (Czech II) genome, we attempted to obtain recombinant clones of restriction fragments containing the common host-virus junction sequences. Tumor DNA B4625 was chosen for this purpose because it appeared to contain only one predominant integrated MMTV (Czech II) genome (Fig. 2).

Tumor B4625 (200 μg) was digested to completion with *Eco*RI and then separated on a preparative agarose gel. A region of the gel containing the MMTV-related 6.7- and 7.5-kb fragments was excised and purified by electroelution (see Materials and Methods). The DNA was cloned into the bacteriophage lambda Wes B vector at the *Eco*RI sites. Recombinant clones (10^5) were screened with the MMTV LTR probe. One positive recombinant clone (designated pA1) was selected for further study. It contains a 7.5-kb *Eco*RI fragment encompassing the 3′ end of the integrated MMTV genome, as well as approximately 4.6-kb of flanking cellular DNA (Fig. 2B). The host flanking sequences were found to contain a short region which hybridized with highly repetitive cellular DNA (Fig. 2B).

To obtain the 5′ host-virus flanking region and the unaltered cellular homolog, tumor DNA was partially digested with the restriction enzyme *Mbo*I to produce fragments with an average size of 15 kb. The DNA was size selected on an agarose gel, purified by electroelution, and ligated into predigested phage EMBL3 at the *Bam*HI recognition site. Recombinant phage (5×10^6) were screened with a mixture of a unique cellular flanking 2.9-kb *Xba*I-*Eco*RI fragment (designated probe A [Fig. 2B]) from pA1 and the MMTV LTR probe. Two positive recombinant clones, Cz4 and Cz5, were isolated. DNA was isolated from each clone and digested with *Sal*I, which cleaves the phage DNA immediately outside the *Bam*HI site used for cloning. *Sal*I fragments of 12.8 and 13.1 kb were purified from Cz4 and Cz5, respectively and then subcloned into pBR322 at the *Sal*I site. The pCz4 recombinant clone contains 7.9 kb of the MMTV (Czech II) proviral genome, including the 5′ LTR, the *gag* and *pol* genes, and a portion of the *env* gene (Fig. 2B). This clone did not, however, react with probe A. Restriction enzyme analysis provided four lines of evidence which support the conclusion that pCz4 and pA1 contain overlapping portions of the same integrated provirus and that pCz5 represents the unoccupied homolog of the flanking cellular sequences. (i) The pCz4 clone contains a 6.7-kb *Eco*RI fragment which reacts with the MMTV LTR and *gag* probes (Fig. 2A [lanes 1 and 3] and B). (ii) The extreme 3′ end of pCz4 contains *Bam*HI and *Eco*RI restriction sites which are also present at the 5′ end of pA1 (Fig. 2B). (iii) The pCz5 clone does not react with the MMTV genome but does hybridize with probe A. Restriction enzyme analysis showed that pCz5 contains restriction sites which are similar to those in the flanking cellular sequences of both pA1 and pCz4 (Fig. 2B). (iv) A unique 2.4-kb *Eco*RI-*Xba*I fragment from pCz5 (designated probe B) hybridized to a similar fragment 5′ to the viral genome in pCz4 (data not shown).

Since Czech II mice had been housed with other inbred mouse strains, it was important to determine whether MMTV (Czech II) was a result of recent horizontal transmission of a laboratory strain of MMTV or was a unique strain which was present in the founder stock of the Czech II colony. Although the MMTVs of *M. musculus* are all highly related, it is possible to distinguish each laboratory strain by restriction site polymorphisms which exist in their proviral DNAs (5, 11, 12, 29). The MMTV (Czech II) proviral genome is approximately 10 kb long and contains three *Eco*RI sites at 2.7, 5.8, and 7.2 kb from the end of the 5′ LTR. *Pst*I sites are found in the LTRs at 200 kb from the 5′ boundary of the LTR. Two additional *Pst*I sites are located at 1.5 and 2.6 kb in the *gag* gene. The proviral genomes of the C3H, GR, and RIII strains of MMTV each lack the *Eco*RI site at 2.7 map units but contain a *Pst*I site at 6.6 map units which is missing in the Czech II viral genome. From these results we concluded that MMTV (Czech II) represents a new strain of MMTV

A common integration region for MMTV (Czech II) in mammary tumor DNA. Comigration of host-virus junction restriction fragments in several tumor DNAs suggested that some Czech II mammary tumors contained an integrated MMTV genome at a common site in the cellular genome (Fig. 1). To confirm and extend these findings, we examined tumor DNAs with the unique cellular flanking sequence probes A and B from tumor B4625. Insertion of proviral DNA into the host genome led to the appearance of novel restriction fragments. Figure 3 shows the results of blot hybridization of *Eco*RI-digested Czech II mammary tumor DNAs with probes A and B. These probes detected the unoccupied 8.6-kb *Eco*RI fragment in liver DNA. We observed no evidence of *Eco*RI restriction fragment length polymorphisms with these probes in the Czech II mouse population. In addition to the 8.6-kb fragment, probes A and B hybridized to other novel fragments in certain tumors. Tumor DNAs N402, B4625, and B4571 each contain a 7.5-kb probe A-related fragment and a 6.7-kb probe B-related fragment. Tumor DNAs B4817 and B4637 each contain unique 8.0- and 6.5-kb fragments which hybridized with probes A and B, respectively. In addition, these probes each hybridized to a 10-kb fragment in tumor B4637. The hybridization signal with the 6.5-, 8.0-, and 10-kb fragments was less intense than with the unaltered 8.6-kb fragment, suggesting that in these tumors there may be cellular heterogeneity with respect to viral insertions in this region of the cellular genome.

To demonstrate that the novel *Eco*RI fragments are the result of MMTV (Czech II) proviral integration, blots were sequentially hybridized with probe B, MMTV LTR, and *gag* probes (Fig. 4). In tumor DNAs N402, B4625, and B4571, the 6.7-kb probe B-related *Eco*RI fragment also reacted with the MMTV LTR and *gag* probes. Similarly, the 6.5-kb fragments in tumor DNAs B4817 and B4637 also reacted with these viral probes. Complementary results were obtained with probe A and the MMTV LTR and *env* gene probes (data not shown). The 10-kb fragment in tumor DNA B4637 reacted with the MMTV LTR but not with the *gag*, *pol*, or *env* gene probes (Fig. 4; data not shown). These results are diagrammatically represented in Fig. 5 and suggest the following. (i) MMTV (Czech II) has integrated within a 0.5-kb region of the cellular genome in five indepen-

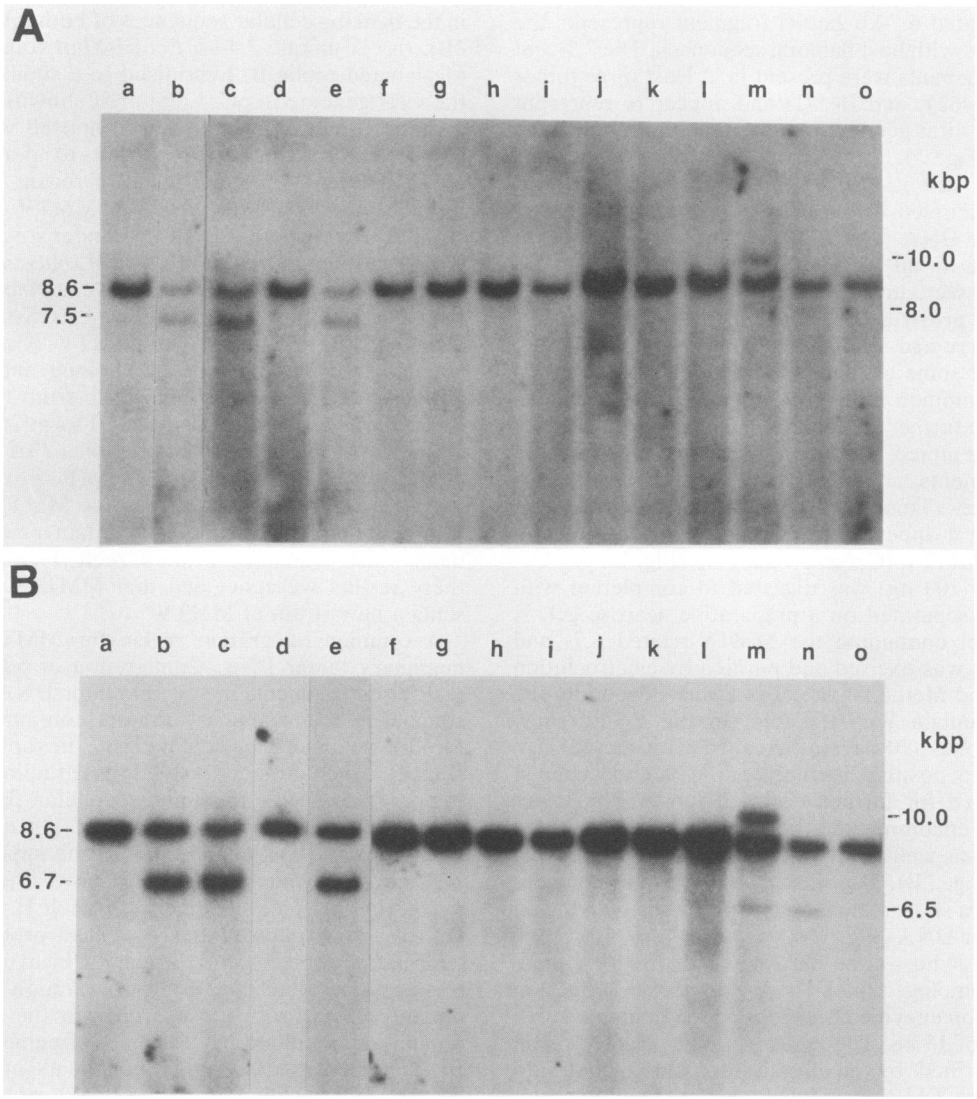

FIG. 3. Rearrangement of *int-3* in mammary tumor cell DNAs. Mammary tumor cell DNAs (10 μg each) were digested with *Eco*RI, run on a 0.8% agarose gel, and transferred to a nylon membrane. The cellular DNAs were from mammary tumors (lanes) B4501 (a), N402 (b), B4625 (c), B4532 (d), B4571 (e), B4579 (f), B4651 (g), B4649 (h), B4661 (i), B4603 (j), B4569 (k), B4681 (l), B4637 (m), B4817 (n), and B4879 no. 2 (o). The membranes were hybridized with (A) probe A or (B) probe B (Fig. 2).

dent tumors, (ii) the transcriptional orientation of each of the complete viral genomes is in the same direction, and (iii) the 10-kb *Eco*RI fragment contains a viral genome in which all but the MMTV LTR sequences have been deleted. Based on these results, we designated this region of the cellular genome *int-3*.

To begin to assess the size of the *int-3* locus, we hybridized DNAs with probes D and E (Fig. 2 and 5). Probe D is a 400-base-pair fragment from clone pCz4 that reacted with a 9.6-kb *Eco*RI fragment in Czech II liver DNA (data not shown). Probe E corresponds to a 4.3-kb *Eco*RI-*Sal*I fragment from pCz5 which lies to the right of the previously identified viral insertion sites in the tumor DNAs. Analysis of *Eco*RI-digested tumor DNAs with this probe revealed a 5.2-kb germ line fragment. Blot hybridization of *Eco*RI-restricted tumor DNAs with probe D or E provided no evidence for additional integrated MMTV genomes within 23.4 kb of the cellular genome (data not shown).

Mammary tumors contain RNA transcribed from *int-3* probe C. MMTV insertion into the *int-1* and *int-2* loci activates transcription of an adjacent cellular gene (8, 21). In these two loci, MMTV integration occurs in clusters which are at either end of the cellular transcribed region. At the 5′ end of the transcribed region, the orientation of MMTV transcription is in the opposite direction from that of the cellular transcribed region. The viral genomes inserted at the 3′ end of the cellular transcribed region are transcribed in the same direction. However, within the MMTV (Czech II) insertion region, all of the viral genomes are oriented in the same direction. Thus, by analogy with the *int-1* and *int-2* loci, we chose to use probes B and C (Fig. 2 and 5), which are unique cellular DNA sequences located 5′ to the MMTV (Czech II) insertion sites. Figure 6 shows a Northern blot of poly(A)-selected RNAs from a lactating mammary gland, an *int-3*-positive tumor (B4625), and an *int-3*-negative tumor (B4661). In Fig. 6B, the RNAs are hybridized with probe C, a cellular flanking *Xba*I-*Bam*HI fragment. This fragment detects a 2.4-kb species of RNA not present in the *int-3*-negative tumor or lactating mammary gland. The same blot was rehybridized with a β-actin probe as a control for the

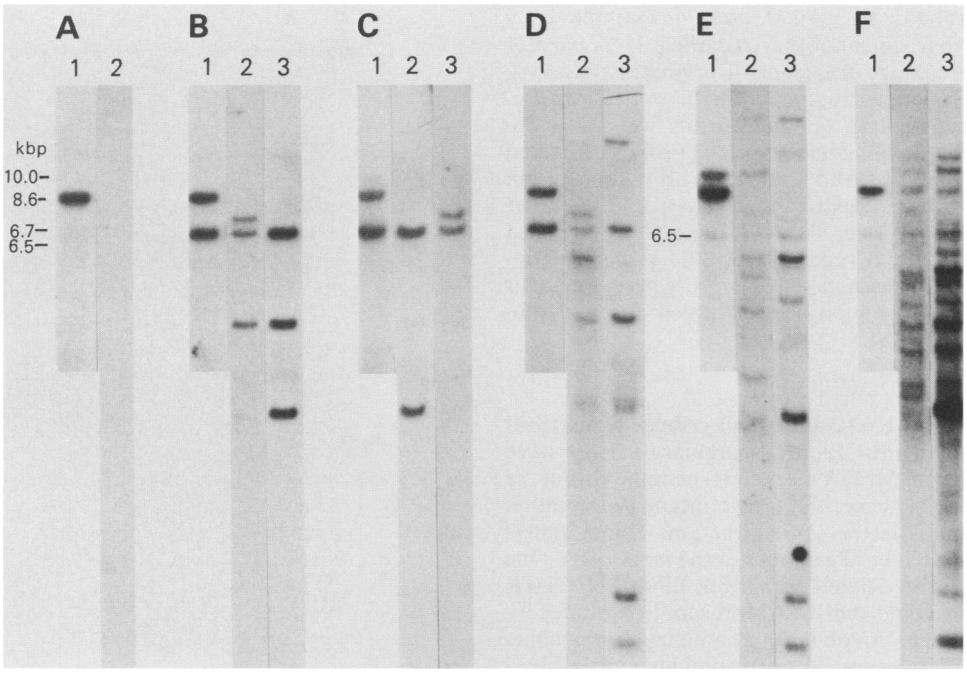

FIG. 4. Correlation between *int-3* rearrangements and MMTV (Czech II)-related sequences in mammary tumor cell DNAs. Normal liver and mammary tumor cellular DNAs (10 μg each) were digested with *Eco*RI, run on a 0.8% agarose gel, and transferred to a nylon membrane. The cellular DNAs were from (panels) normal liver (A), tumor B4625 (B), tumor N402 (C), tumor B4571 (D), tumor B4637 (E), and tumor (B4817) (F). The membranes were sequentially hybridized with probe B (lanes 1), MMTV LTR (lanes 2), and MMTV *gag* (lanes 3) probes (see the text).

amount of RNA present. The 2.2-kb β-actin RNA species was detected in each sample. Another Czech II *int-3*-positive tumor (B4571) was also found to contain the 2.4-kb cellular RNA species detected with probe C (data not shown). Thus, the *int-3* transcribed region lies 5′ to the sites where MMTV (Czech II) had been inserted into the cellular genome. Since probe B did not hybridize to these RNA samples (data not shown), these sequences may correspond to an intron within the *int-3* transcribed region.

The *int-3* locus contains a highly conserved cellular gene. There are approximately 40 cellular proto-oncogenes whose expression, when activated, could contribute to tumor development (2). To determine whether *int-3* corresponds to any of these genes, we began to determine whether it shares nucleotide sequence homology with representatives of known proto-oncogenes (c-*erb* B, N-*ras*, cH-*ras*-1, cK-*ras*-2, N-*myc*, c-*myc*, and c-*mos*), viral oncogenes (v-*abl*, v-*src*, v-*fms*, v-*fes*, and v-*sis*), or *int-1* and *int-2*. Under stringent blot hybridization conditions, *int-3* probe C did not react with any of the oncogenes, including *int-1* and *int-2* (data not shown).

One characteristic of proto-oncogenes, as well as the *int-1* and *int-2* loci, is that they have been evolutionarily conserved (2, 4, 20). The evolutionary conservation of the *int-3* locus was examined in cellular DNAs from different species, including *Salmo* sp. (salmon sperm), *M. musculus* subsp. *domesticus* (C3H), *M. musculus* subsp. *musculus* (Czech II), *M. cervicolor* subsp. *cervicolor*, *Gallus gallus* (chicken), *Equus caballus* (horse), and *Homo sapiens* (human). These DNAs were digested with *Eco*RI and blot hybridized with probe C under stringent conditions (Fig. 7). DNAs from all of the mammalian species examined hybridized under these

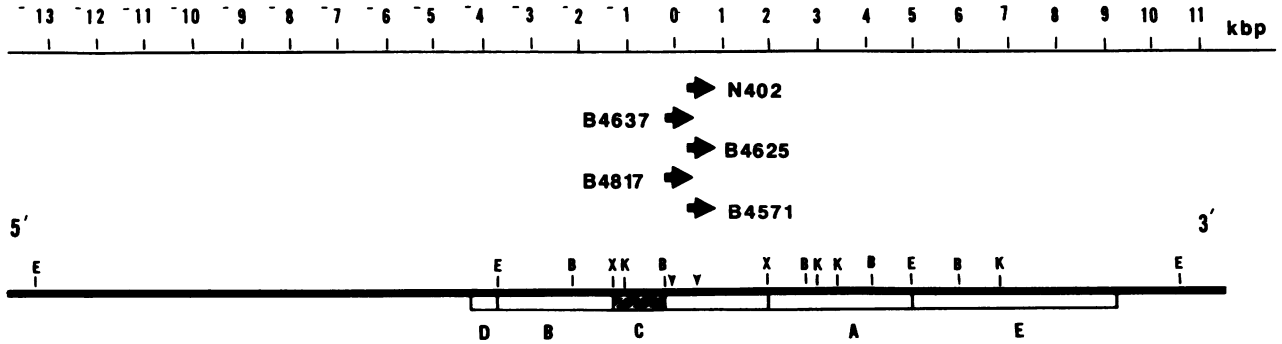

FIG. 5. Restriction map of the *int-3* locus. The restriction sites are indicated as follows: E, *Eco*RI; B, *Bam*HI; K, *Kpn*I; X, *Xba*I. The sites of the two clusters of integrated MMTV (Czech II) genomes are indicated by the two vertical arrowheads. The transcriptional orientation of the integrated viral genomes is indicated by the horizontal arrowheads. The origins of the cellular flanking sequence probes are indicated by open boxes. The closed and hatched box corresponds to probe C, which hybridized with a 2.4-kb species of tumor RNA.

conditions with probe C. The two *M. musculus* samples each contained an 8.6-kb fragment observed earlier. *M. cervicolor* possessed a different-size fragment of approximately 20.0 kb, indicating restriction fragment length polymorphism. Restriction fragment length polymorphisms were also observed in the other mammalian species. Probe C reacted weakly with chicken DNA and not at all with salmon sperm DNA. Probe A did not react with DNA from any other mammalian species (data not shown). These results suggest that the *int-3* transcribed region (probe C) is highly conserved, at least among mammalian species, whereas cellular sequences flanking this region have diverged during evolution.

DISCUSSION

The topography of the *int-3* locus and cellular transcribed region. Five independent Czech II mammary tumors have been shown to have an MMTV (Czech II) genome within the *int-3* locus. In each of these, the transcriptional orientation of the viral genome is in the same direction. Tumor B4637 contains two integrated viral genomes at the *int-3* locus. One of these appears to be deleted for all but the MMTV LTR sequences. It is possible that this represents two independent integration events involving a complete viral genome and a deleted proviral genome corresponding to an aberrantly spliced MMTV RNA. A more plausible possibility is that the deleted viral genome arose from the complete viral genome during evolution of the tumor. This could have resulted from a recombination event between the two MMTV LTR elements, which would cause loss of the intervening structural genes. Dickson et al. (8) also have described a tumor with a single MMTV LTR present in the *int-2* locus. Further restriction mapping of recombinant DNA clones containing the two proviral genomes would be necessary to answer this question in an unambiguous manner.

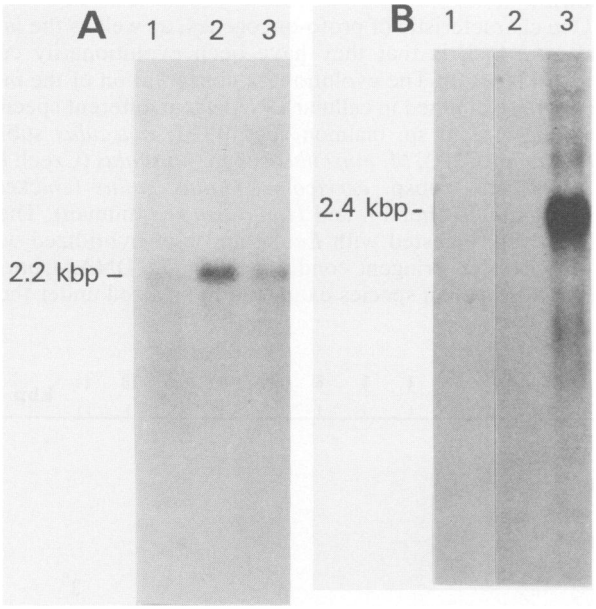

FIG. 6. Expression of *int-3* in *M. musculus* subsp. musculus (Czech II) mammary tumors. Poly(A)$^+$ RNAs (3 μg each) prepared (see the text) from normal lactating mammary gland (lane 1), the *int-3*-negative tumor B4661 (lane 2), and the *int-3*-positive tumor B4625 (lane 3) were denatured in the presence of formaldehyde and run on 1% agarose gels containing formaldehyde. The RNA samples were then transferred to a nylon membrane and hybridized with either β-actin probe (panel A) or *int-3* probe C (panel B).

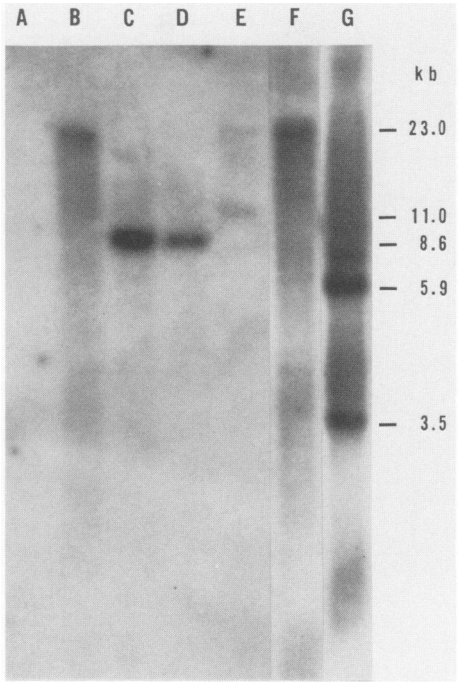

FIG. 7. Sequences related to the *int-3* locus in other vertebrate species. Cellular DNAs (10 μg each) from (lanes) salmon sperm (A), *M. cervicolor* subsp. *cervicolor* (B), *M. musculus* subsp. *musculus* (C), *M. musculus* subsp. *domesticus* (D), horse (E), chicken (F), and human (G), were digested with *Eco*RI, run on a 0.8% agarose gel, and transferred to a nylon membrane. The blot was hybridized under stringent conditions (see the text) with probe C.

MMTV insertion into the *int-3* locus activates expression of a 2.4-kb species of RNA which is not present in lactating mammary tissue or in a tumor lacking a viral insertion at the *int-3* locus. However, this RNA species was not detected with probe B, which corresponds to 2.9 kb of unique cellular DNA sequences adjacent to the 5′ end of probe C. This suggests that the *int-3* transcribed region contains a minimum of two exons which are separated by an intron that is 2.9 kb or more in size. Since no further analysis of the *int-3* transcript was attempted, the precise organization of introns and exons and the transcriptional orientation of the cellular transcribed region remains to be determined. However, it is possible to make certain predictions by comparing the organization of the *int-3* locus with that of the *int-1* and *int-2* loci. The cellular transcribed regions of the *int-1* and *int-2* loci are, respectively, 5.0 and 7.0 kb in size. In each case, the cellular transcribed sequences are bounded by regions of tumor DNA which are frequently occupied by an MMTV genome (8, 21). The transcriptional orientation of the viral genomes which are integrated at the 5′ end of these loci is in the opposite direction of that of the cellular transcribed region, whereas transcription of the viral genomes occurring at the 3′ end of the locus is in the same direction as the cellular transcribed region. By analogy with these loci, the cellular transcribed region of *int-3* is likely to be much larger. This is suggested by the absence of viral integration events within 13.0 kb of tumor DNA that lies 5′ to the previously identified cluster of viral insertions at *int-3*. Thus, it is possible that further examination of DNA sequences more distant to the 5′ end of those already examined will reveal

another region of tumor DNA which is frequently occupied by an MMTV (Czech II) genome. This would represent the other boundary of the *int-3* transcribed region.

Expression of *int-3* RNA does not begin within the LTR of the MMTV (Czech II) proviral genome since the 2.4-kb *int-3* RNA does not react with the LTR probe (data not shown). There are at least two possible mechanisms by which MMTV (Czech II) activates *int-3* RNA expression. First, integration of an MMTV (Czech II) genome at *int-3* could disrupt *cis*-acting regulatory sequences which control RNA expression at this locus. Another possibility which would explain the uniform transcriptional orientation of the proviral genomes with respect to the *int-3* transcribed region has been proposed by Nusse and Varmus (21) and Dickson et al. (8). In their model, enhancer elements within the U3 region of the MMTV LTR are responsible for activation of the adjacent cellular transcribed region. These elements could act only or preferentially on cellular transcriptional promoters which are 5′ to the integrated proviral genome.

Possible role in tumorigenesis. At present, there are primarily two lines of circumstantial evidence supporting the view that MMTV-activated cellular genes contribute to mammary tumor development. First, there is a high frequency of mammary tumors in which expression of one of the *int* loci is activated by an MMTV genome integrated in adjacent cellular sequences. Second is the analogy of this phenomenon with other retrovirally induced neoplastic diseases. Indeed, integration of a retroviral genome at a common integration locus in several tumors of the same type has been reported in avian B-cell lymphomas induced by two unrelated type C retroviruses (15, 19), avian erythroleukemia induced by a type C retrovirus (13), and T-cell lymphomas induced by murine leukemia virus (6, 34). In each case, the consequence of viral integration at the common loci in tumor DNA is activation of expression of an adjacent cellular gene. The common integration loci in the B-cell lymphomas and avian erythroleukemia are, respectively, the c-*myc* and c-*erb* B proto-oncogenes. These genes are the progenitors of the acute tumorigenic retroviral oncogenes. Most of the common integration regions in T-cell tumors, like the MMTV *int* loci, appear to be unrelated to the known proto-oncogenes. However, each of these cellular genes, like the proto-oncogenes, are highly conserved among mammalian and avian species.

The biology of MMTV-induced mammary tumors appears to involve a multistep process (37). In MMTV (C3H)- and MMTV (Czech II)-infected mice, the earliest detectable stage of tumor development is the appearance of hyperplastic alveolar cells or nodules. Hyperplastic alveolar nodules are considered premalignant precursors to the later developing pregnancy-independent adenocarcinomas. The MMTV (RIII) strain induces pregnancy-dependent mammary tumors which, after two or more pregnancies, progress to hormone independence. It seems probable, therefore, that multiple MMTV-induced or stochastic mutations are required for tumor development. Indeed, activation of *int-1* and *int-3* is associated with development of virally induced, pregnancy-independent mammary adenocarcinomas. Activation of the *int-2* locus occurs very early in the pregnancy-dependent stage of MMTV (RIII)-induced mammary tumors (24). Additional mutations are probably required for progression to hormone independence. Peters et al. (25) have provided evidence consistent with this conclusion. They found, in 30 MMTV (RIII)-induced mammary tumors, 15 with viral insertions in both the *int-1* and *int-2* loci and six each with an insertion at *int-1* or *int-2* only. Recently, they have shown that two of the three remaining tumors that lack viral insertions at *int-1* or *int-2* contain a viral insertion at *int-3* (G. Peters, personal communication). The necessity for multiple genetic alterations during mammary tumor development has also been suggested for transgenic strains of mice containing a recombinant c-*myc* gene linked to the MMTV LTR (32). Although these mice have a high incidence of mammary adenocarcinomas, not every mammary cell is tumorigenic, and tumors appear only after two or more pregnancies. These observations suggest to us that the *int* loci represent a family of cellular genes which, when activated by MMTV, contribute to deregulation of mammary gland development.

The different MMTV strains isolated from inbred and feral *M. musculus* are all highly related and appear capable of integrating at many sites within the cellular genome. Thus, it seems a paradox that the frequency with which the different *int* loci are activated varies with the particular strain of MMTV (10, 21, 25). Since different strains of mice were used in these studies, it is not possible to evaluate the potential contribution of the genetic background of the host to the apparent specificity. It is known, however, that the kind of tumor (pregnancy dependent or independent) induced by MMTV is associated with the particular strain of virus and is independent of the mouse strain (31). Similarly, it could be that the relative frequencies with which different *int* loci are activated reflect a property of the virus which distinguishes different target cell populations in developing mammary glands. It will be of interest to determine whether mammary tumors induced by new strains of MMTV from feral mice (10; Callahan, unpublished data) reveal additional *int* loci in tumor DNA.

ACKNOWLEDGMENTS

We thank E. Hogg and J. Robbins for expert technical assistance. We are indebted to M. Potter for many helpful discussions and encouragement during the course of these experiments. Several colleagues kindly provided recombinant DNA clones of MMTV proviral DNA (H. Varmus, J. Majors, and N. Hynes), v-*fes* and v *fms* (C. Sherr), c-*sis* (R. Gallo), c-Ha-*ras*-1 (E. Chang), β-actin (J. Featherston), *int*-1 (R. Nusse), and *int-2* (C. Dickson and G. Peters).

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Bishop, J.** 1985. Viral oncogenes. Cell **42**:23–38.
3. **Callahan, R., W. Drohan, D. Gallahan, L. D'Hoostelaere, and M. Potter.** 1982. Novel class of mouse mammary tumor virus related DNA sequences found in all species of *Mus*, including mice lacking the virus proviral genome. Proc. Natl. Acad. Sci. USA **79**:4113–4117.
4. **Casey, G., R. Smith, D. McGillivray, G. Peters, and C. Dickson.** 1986. Characterization and chromosome assignment of the human homolog of *int-2*, a potential proto-oncogene. Mol. Cell. Biol. **6**:502–510.
5. **Cohen, J. C., P. R. Shank, V. L. Morris, R. Cardiff, and H. Varmus.** 1979. Integration of the DNA of mouse mammary tumor virus in virus-infected normal and neoplastic tissue of the mouse. Cell **16**:333–345.
6. **Cuypers, H. T., G. Selten, W. Quint, M. Zijstra, B. R. Maandag, W. Boelens, P. Van Wezenbeek, C. Melief, and A. Berns.** 1984. Murine leukemia virus induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. Cell **37**:141–150.
7. **Denhardt, P. T.** 1966. A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641–646.
8. **Dickson, C., R. Smith, S. Brookes, and G. Peters.** 1984. Tumorigenesis by mouse mammary tumor virus: proviral acti-

vation of a cellular gene in the common integration region *int*-2. Cell **37**:529–536.
9. **Dunn, T. B.** 1958. Morphology of mammary tumors in mice, p. 30–84. *In* F. Homburger (ed.), The physiopathology of cancer, 2nd ed. Hoeber-Harper, New York.
10. **Escot, C., E. Hogg, and R. Callahan.** 1986. Mammary tumorigenesis in feral *Mus cervicolor popaeus*. J. Virol. **58**:619–625.
11. **Etkind, P. R., P. Szabo, and N. H. Sarkar.** 1982. Restriction endonuclease mapping of the proviral DNA of the exogenous RIII murine mammary tumor virus. J. Virol. **41**:855–867.
12. **Fanning, T. G., J. P. Puma, and R. D. Cardiff.** 1980. Selective amplification of mouse mammary tumor virus in mammary tumors of GR mice. J. Virol. **36**:109–114.
13. **Fung, Y. K., W. G. Lewis, L. B. Crittenden, and H. J. Kung.** 1983. Activation of the cellular oncogene c-*erb* B by LTR insertion: molecular basis for induction of erythroblastosis by avian leukosis virus. Cell **33**:357–368.
14. **Gross-Bellard, M., P. Oudet, and P. Chambon.** 1973. Isolation of high molecular weight DNA from mammalian cells. Eur. J. Biochem. **36**:32–38.
15. **Hayward, W., B. Neel, and S. Astrin.** 1981. Activation of a cellular *onc* gene by promotion, insertion in ALV-induced lymphoid leukosis. Nature (London) **290**:475–480.
16. **Maniatis, T., B. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Moore, R., G. Casey, S. Brookes, M. Dixon, G. Peters, and C. Dickson.** 1986. Sequence, topography and protein coding potential of mouse *int*-2: a putative oncogene activated by mouse mammary tumor virus. EMBO J. **5**:919–924.
18. **Morris, V., E. Medeiros, G. Ringold, J. Bishop, and H. Varmus.** 1977. Comparison of mouse mammary tumor virus specific DNA in inbred and Asia mice and in tumors and normal organs from inbred mice. J. Mol. Biol. **114**:73–91.
19. **Noori-Daloii, M. R., R. A. Swift, N. J. Kung, L. B. Crittenden, and R. L. Witter.** 1981. Specific integration of REV proviruses in avian bursal lymphomas. Nature (London) **294**:574–576.
20. **Nusse, R., A. van Ooyen, D. Cox, Y.-K. Fung, and H. Varmus.** 1984. Mode of proviral activation of a putative mammary oncogene (*int*-1) on mouse chromosome 15. Nature (London) **307**:131–136.
21. **Nusse, R., and H. Varmus.** 1982. Mammary tumor induced by the mouse mammary tumor virus: evidence for a common region for provirus integration in the same region of the host genome. Cell **31**:99–109.
22. **Peters, G., S. Brookes, R. Smith, and C. Dickson.** 1983. Tumorigenesis by mouse mammary tumor virus: evidence for a common region for provirus integration in mammary tumors. Cell **33**:369–377.
23. **Peters, G., C. Kozak, and C. Dickson.** 1984. Mouse mammary tumor virus integration regions *int-1* and *int-2* map different mouse chromosomes. Mol. Cell. Biol. **4**:375–378.
24. **Peters, G., A. E. Lee, and C. Dickson.** 1984. Activation of cellular gene by mouse mammary tumor virus may occur early in mammary tumor development. Nature (London) **309**: 273–275.
25. **Peters, G., A. E. Lee, and C. Dickson.** 1986. Concerted activation of two potential proto-oncogenes in carcinomas induced by mouse mammary tumor virus. Nature (London) **320**:628–631.
26. **Rigby, P. W. J., M. Dieckman, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
27. **Robbins, J. M., D. Gallahan, E. Hogg, C. Kozak, and R. Callahan.** 1986. An endogenous mouse mammary tumor virus genome common in inbred mouse strains is located on chromosome 6. J. Virol. **57**:709–713.
28. **Schlom, J., W. Drohan, Y. Teramoto, P. Hand, D. Colcher, R. Callahan, G. Todaro, D. Kufe, D. Howard, J. Gautsch, R. Lerner, and G. Schidlovsky.** 1978. Diversity of mouse mammary tumor virus genetic information and gene products in rodents, p. 342–368. *In* H. Morris (ed.), Origins of inbred mice. Academic Press, Inc., New York.
29. **Shank, P. R., J. C. Cohen, H. E. Varmus, K. R. Yamamoto, and G. M. Ringold.** 1978. Mapping of linear and circular forms of mouse mammary tumor virus DNA with restriction endonucleases: evidence for a large specific deletion occurring at high frequency during circularization. Proc. Natl. Acad. Sci. USA **75**:2112–2116.
30. **Southern, R. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
31. **Squartini, F., G. Rossi, and J. Paoletti.** 1963. Characters of mammary tumors in BALB/c female mice foster nursed by C3H and RIII mothers. Nature (London) **197**:505–506.
32. **Stewart, T., P. Pattengale, and P. Leder.** 1984. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/*myc* fusion genes. Cell **38**:627–637.
33. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
34. **Tsichlis, P., P. G. Strauss, and L. F. Hu.** 1983. A common region for proviral DNA integration in MoMuLV-induced rat thymic lymphomas. Nature (London) **302**:445–449.
35. **van Ooyen, A., and R. Nusse.** 1984. Structure and nucleotide sequence of the putative mammary oncogene *int*-1: proviral insertions leave the protein-encoding domain intact. Cell **39**:233–240.
36. **Varmus, H. E.** 1982. Recent evidence for oncogenesis by insertion mutagenesis and gene activation. Cancer Surv. **1**:309–319.
37. **Weiss, R. N., N. Teich, H. Varmus, and J. Coffin.** 1982. Origins of contemporary RNA virus research, p. 16–20. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), Molecular biology of tumor viruses: RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.